Stahl, S.M. and Wets, K. (1987). "Indoleamines and schizophrenia". In: Handbook of Schizophrenia Vol. # 2: Neurochemistry and Neuropharmacology of Schizophrenia, (eds F.A. Henn and L.E. DeLisi), Elsevier Science Publishers, Amsterdam, pp. 257-296.

Van Tol, H.H.M., Bunzow, J.R., Guan, H.C., et al. (1991). "Cloning of the gene for a human dopamine D_4 receptor with affinity for the antipsychotic clozapine", Nature, 350, 610 - 614.

9

AN EXAMINATION OF THE COMPLEXITIES IN THE MEASUREMENT OF RECOVERY IN SEVERE PSYCHIATRIC DISORDERS

Courtenay M. Harding

Results from five recent investigations studying the very long-term"outcome" of schizophrenia (in Switzerland, Germany, and the United States) indicate that no matter how chronic the cohorts were, approximately 25% of the subjects achieved recovery at follow-up, with an additional 25 - 40% improved, with achievement of wide heterogeneity as the rule rather than the exception (Bleuler, 1972; Ciompi and Moller, 1976; Harding et al., 1987b, Huber et al., 1979; Tsuang et al., 1979). In fact, these studies and other shorter ones (Bland and Orn, 1980; Gardos et al., 1982; Hawk et al., 1975, Strauss and Carpenter, 1974; Strauss et al., 1978; WHO, 1979) have shown that the course of severe psychiatric disorder is a complex, dynamic, and heterogeneous process, which is non-linear in its patterns moving toward significant improvement over time and helped along by an active, developing person in interaction with his or her environment (Harding and Strauss, 1985).

Our new appreciation about this process evokes many questions. For instance, just what is meant by the use of the word "recovery"? Is the subject simply no longer symptomatic? Is she or he considered to be "recovered" or just "in remission"? Has the person re-found (or recovered) all his or her initial levels of functioning which were lost during a psychotic episode or do we mean that he or she is functioning more soundly and stronger than ever before? Against which

Schizophrenia: Exploring the Spectrum of Psychosis. Edited by R. Ancill.

baseline do we make such judgments? Is a new episode after "recovery" considered to be a new illness? What do we mean when we discuss "outcome"?

After the completion of two three-decade studies of the long-distance course of schizophrenia, the author examines some of these questions which arise in the measurement of recovery from severe psychiatric disorder. She suggests that constructs such as "outcome" and "end state" often lead to inaccurate conclusions about disorders. In contrast to cross-sectional measures, longitudinal assessments are portrayed as essential to an accurate understanding of prolonged illness.

To cope with these challenges, puzzled investigators have employed four primary strategies to determine the amount of recovery and improvement. The strategies comprise the following: (1) measurement against each subject's own baseline, (2) employment of grouped means, (3) comparison across studies, and (4) assessment against a construct of "normal" behavior. This chapter will discuss these strategies and present additional suggestions proposed by the Vermont and Maine Longitudinal Research Project teams as we attempted to undertake this prodigious task ourselves. We have organized our thinking along five major themes and propose that: (1) the concept of outcome is a research artifact, (2) "two cross-sectionals do not a longitudinal make", (3) global measures of psychopathology and recovery are not sufficient, (4) double standards are often employed in assessments, and (5) a rater's epistemological framework about illness, recovery, and human beings plays an important role in measurement in that "what you are trained to ask about and observe, strongly determines what you ask about and observe". The author concludes with the suggestion that these factors do indeed color the judgment of which patients are considered to be recovered and which are not.

AN OVERVIEW OF THE VERMONT AND MAINE LONGITUDINAL STUDIES OF SEVERE PSYCHIATRIC DISORDER

In the Vermont Project, 269 Vermont State Hospital patients with severe and persistent psychiatric disorders (especially with schizophrenia) were originally selected for very chronic disability, rehabilitated in a model demonstration program, and released to the community in the mid-1950s (Chittick et al., 1961). At the time of entry into this project in the 1950's, these subjects had an average of six years of continuous psychiatric hospitalization and sixteen years of disability. A five to ten-year follow-up study in the 1960s revealed that two thirds of the subjects were in the community maintained by considerable investment of time, money, and effort (Deane and Brooks, 1967).

Twenty to twenty-five years after their entry into this program, 97% of this group were located and assessed. The catamnestic period averaged 32 years. Two field interviews were conducted by raters blind to record information. The first interview provided a multivariate cross-sectional assessment of outcome. The second interview included a Meyerian/Leighton Life Chart (Leighton and Leighton, 1949; Meyer, 1919) and contributed a longitudinal documentation of patterns, shifts, and trends in the course of life for members of the cohort. These patterns were derived from a multidimensional year-by-year documentation of the 20 to 25-year period. Close relatives and others, who knew the subjects well, were interviewed to verify current status and historical data for both the live and deceased subjects. New and traditional scales were used to provide structured measures of outcome. Hospital and Vocational Rehabilitation records were independently reviewed with the Hospital Record Review Form battery (HRRF), compiled from a modification of WHO's Psychiatric and Personal History form (WHO, 1979) and Strauss's Case Record Rating Scale (Strauss and Harder, 1981). The record raters were blind to outcome and field data. All batteries were subjected to inter-rater trials and inter-item concordance testing (Harding et al., 1987a; 1987b).

In order to make our subjects comparable to today's patients we then, applied the DSM-III criteria (APA, 1980) to their index status with records that were stripped of all previous diagnostic assignments (Harding et al., 1987b). John Strauss at Yale and Alan Breier at NIMH completed the rediagnostic work after two sets of inter-rater trials. Analysis of the long-term outcome for those subjects, who once qualified for DSM-III schizophrenia, revealed wide heterogeneity. For one-half to two-thirds of the group, course was neither downward nor marginal. This finding was remarkable because these patients represented those in the "bottom third" of the schizophrenia spectrum and the back wards in the hospital. Most people resided in the community, were able to care for themselves, had become actively involved with family and friends, and made productive contributions to their families and communities with little or no residual display of symptomatology (Harding et al., 1987b). These findings mirrored those derived for the full cohort, which included patients with a wide variety of other DSM-III diagnoses (Harding et al., 1987a). Thus, the more stringent criteria of DSM-III were not associated with uniformly poor outcome as expected (APA, 1980). The definition of recovery included the following factors measured at follow-up: no signs or symptoms of psychiatric disorder, no psychotropic medications, working or retired appropriately after a work history (especially important because the average age at follow-up was 61 ranging up to 79 years of age), maintaining mutually satisfying interdependent relationships, the absence of behavioral or contextual indicators that they were former mental

patients, and integration into the community as full-fledged citizens. We agreed with Vaillant's conclusion after his ten-year study, that "diagnosis and prognosis should be treated as two different dimensions of psychosis" (Vaillant, 1975).

The Vermonters received a ten-year pioneering and comprehensive psychosocial and vocational rehabilitation program. We were unable to determine if the program had made an impact upon the outcome findings because all the patients, who were in the back wards of the state's only state hospital, were selected with the exception of those on legal mandates, developmental disabilities, or over 62 years of age. However, Augusta State Hospital in Maine was another New England hospital with a similar catchment area and hospital which did not give a rehabilitation program to its patients. We achieved a comparison sample, person by person computer-matched to the Vermont cohort members on age, sex, diagnosis, and length of hospitalization during the window of 1956-61. The Maine cohort was then followed with the same protocol, the same diagnostic criteria, and instrument batteries with both intra- and inter-project reliabilities established as well as blindness in the data collection. Ninety-four per cent of the Mainers were assessed at an average of 36 catamnestic years. This event was the first time in longitudinal research that two such long-term studies had matched samples, protocols, diagnostic criteria, and historical periods (DeSisto et al., submitted, a).

The Mainers did significantly less well than did the Vermonters both in the quantity and quality of recovery, especially in areas of work, symptoms, global outcome (even covarying out other significant modifiers (e.g. acute onset, education, urban/rural settings). Longitudinal year by year patterns also showed more positive and independent trajectories for Vermonters. We concluded that rehabilitation and the opportunity to be out of the hospital joined with biological correction mechanisms to potentiate a return to the highest level of function possible for each person (DeSisto et al., submitted, b).

THE MEASUREMENT OF RECOVERY

Outcome is a Research Artifact

The Vermont and Maine Projects are considered to be studies about the long-term outcome of schizophrenia. They represent some of the longest studies in the world literature, with catamneses of 22 to 62 years after first admission as well as projects with the most stringent methodology. Other studies have looked at 5, 10, 20, and 25 years post admission and each of those projects is also called an outcome study. I propose that each one of those studies is a research artifact;

that one investigator's "outcome" level is another investigator's "course". For example, those projects which have completed a 10-year follow-up consider a 5-year marker as part of the course; a 20-year follow-up considers both the 10-year and the 5-year points in time as part of the course of the 20 years (Harding and Strauss, 1985). The Vermont marker at 62 years has been superseded by Luc Ciompi's study at 64 years (Ciompi, 1980), ad infinitum, to its own logical conclusion. Therefore, we suggest that instead of assessing outcome we are actually studying markers in the course of life (Harding and Strauss, 1985) in which illness is only a part.

The largest misnomer to date is the so-called "end state", used by Manfred Bleuler (1978) and other Europeans to describe plateau states of five years in which subjects have achieved a certain level of recovery and a stabilized period. Use of the words "end state" gives the incorrect impression that a person will stay at that state and that it is the final state. In fact, Bleuler (1978) did not intend to imply such a finality (pp. 190 - 192). Bleuler, himself, said that there were sometimes late improvements after as many as 30 to 40 years of illness (Bleuler, 1978, pp. 228 - 233). In reality, the words usually denote an end state for the research project itself. Although an investigator's device, the phenomenon might be re-labeled and described from a different vantage point. The knowledge of plateaus and late improvements might be valuable contributions to the biopsychosocial understanding of the course of schizophrenia and should have a new name.

"Two Cross-Sectionals Do Not a Longitudinal Make"

Many investigations start by achieving a baseline measurement on all subjects, and then a number of years later secure another cross-sectional measurement at that point in time - or vice versa. These studies have been labeled longitudinal, and we have suggested that "two cross-sectionals do not a longitudinal make" (Harding and Strauss, 1985, p. 3421). By way of analogy, a cross-sectional could be likened to a black-and-white portrait. The observer can tell gender, approximate age, something about the surrounding background, but is unable to determine whether the wrinkles around the forehead and the eyes come from a lot of laughter or many frowns. Raters are unable to tell whether that background is a usual one or not; nor are they able to tell how the subject got to that spot where he or she is today. In the cross-sectional part of the interviews with both the Vermont and Maine cohorts, we were aware of the fact that the month we happened to interview people was purely arbitrary. We knew that we had found some people who were having a good month when they ordinarily did not, but we realized that the opposite was also true, and hoped they balanced out.

However, a longitudinal can be a follow-along or a follow-back through a block of time year after year with many measures across many domains. Because human life is very difficult to capture in all its dynamic ebbs and flows, and since it is a longitudinal process, a year-by-year accounting comes much closer to reality. Continuing the earlier analogy, this strategy would be more akin to watching a theater-in-the-round in which one can gain perspective from looking at the same person from different angles over time with characters interacting and changing. Given this distinct bias, I would venture an opinion that few judgments can be made about degrees of recovery or lack of it unless there is a longitudinal documentation of course.

Ways to Assess a Heterogeneous Process

Early investigators, such as E. Bleuler (1950), Holt and Holt (1952), Stephens (1970) and others, used global and often undefined terms of "recovered or "not recovered". Other researchers divided recovery into symptom recovery and social recovery, quite often loosely defined but at least a step in the right direction (e.g., Harris et al., 1956; Holmbe and Astrup, 1957; Lo and Lo, 1977; Rennie, 1939). The question remains, "How do we define and measure recovery?"

The Vermont and Maine Projects chose to triangulate data by using measures of functioning, such as the Global Assessment Scale (GAS) by Endicott et al. (1976); Levels of Function Scale (LOF), by Strauss and Carpenter (LOF) by Hawk, Carpenter and Strauss, (1975); the Brief Psychiatric Rating Scale (BPRS) by Overall and Gorham (1962), as well as 12 other classic scales rating twenty-one domains (Harding et al., 1987a). From the cumulative data acquired over three hours of field interviews with the probands themselves, we subsequently also rated them for adjustment - another guise for "recovery". The Community Adjustment Scale (CAS) was designed by Consalvo et al. (1981) to measure adjustment based on scores for productivity, intimacy, and behavior (see Figure 1 below).

The findings from this scale were significantly related to those generated by the GAS and Level of Function Scale in a chi-square analysis ($p = <.001$ [CAS vs. GAS] and $p = <.004$ [CAS vs LOF)]. A little over one quarter (29%) of the Vermont cohort were rated as functioning well across all domains. Recovery is not difficult to determine or defend with that group. A group of about 36% of the Vermonters, whose functioning across all domains was very poor, also was not difficult to define as "not recovered".

However, there were 35% in Vermont who were in between those two extremes. They presented serious problems for the rater. There were people who still had positive signs of schizophrenia (e.g., hallucinations, and/or delusions) but who were quite functional people. They retained mild impairment, but no disability. They worked, had families, friends, were generally satisfied with life, and had

Figure 1. Community Adjustment Scale (CAS)* (Consalvo et al., 1981)

1. Criteria

1. Degree of productivity based on work (housework, volunteer, avocation), or retired and functioning at a level appropriate to age and health. Rate on a one to five scale.

2. Degree of intimacy achieved, based on the nature of interpersonal relationships (e.g., marriage, friendship, kinship), and the possession of interpersonal skills. Rate on a one to five scale.

3. The relative absence of behavioral (e.g., bizarre speech, actions, appearance), or contextual (e.g., living in a boarding home), indications that they are former mental patients. Rate on a one to five scale.

2. Ratings for overall adjustment Use total points from above three criteria.

1. Well adjusted. Individuals in this group exhibit a moderate to high level of adjustment in all three areas and would be described as without any indications of being a former mental patient. Total range 12 - 15 points.

2. Well adjusted but... Individuals in this group ordinarily have a moderate to high level of adjustment in all areas but one, or a marginally moderate level of adjustment in at least two of the three criteria categories. Such individuals would give no clear indication that they were ever former mental patients but their overall level of adjustment would not rule out that possibility. Total range 8 - 11 points.

3. Maladjusted. These individuals display poor adjustment in at least two criteria areas and could at least be viewed as probable former mental patients. Total point range 3- 7 points.

* It should be noted that this scale was completed after a comprehensive structured interview across 21 areas of functioning and psychological status.

learned to control their symptoms. Would they be labeled recovered? We said they are significantly improved and rated them 'Well, but..." There was another group which was quite sociable, maintained supportive interpersonal relationships, had hobbies, and was quite happy, but did not work. They told us of the disincentives in the entitlement systems and its environmental impact on

their lives. Another group worked well, but was composed of self-described "loners" with no family and no friends with whom to interact. Is this group improved? Many raters would hastily say "No, not even improved." Let us suppose that some of those so-called loners have always been loners and preferred their own company to others? Was their behavior prodromal and now considered to be residual? Or, are there not substantial numbers of people, both outside of or working for the mental health system who live alone. They are quite happily functioning at work and caring for themselves and who do not maintain relationships? Are we not asserting our own conceptions about "normal" behavior?

STRATEGIES TO DETERMINE CURRENT STATUS

Strategy #1: Assessment of Abnormality Against Societal Norms

What then is the range of normal human functioning? Hogarty and Katz (1971) produced some work in the early 1970s assessing 450 non-patients in one Maryland county to acquire norms for their instruments. They discovered that patterns of behavior for age, marital status, social class, and gender must be taken into consideration when making judgments (e.g., normal adolescents were reported to show negativism, general psychopathology, and less stability). They asked: "Is it valid to point in absoluteness to the belligerence, negativism, and poor performance of juvenile offenders and otherwise 'disturbed adolescents' when corresponding norms of age-related behavior are so characterized?" (Hogarty and Katz, 1971, p.479). It is a very important question.

As another example, the difference in gender functioning has only recently become appreciated. Holstein and Harding (1992) have assessed the data for women only from the Yale Longitudinal Study in which we were following people intensively for the first two years after episode (Strauss et al., 1985). We found that those women who were rated as more symptomatic were often those who carried dual work roles and were quite functional in caring for home and family, as well as working. These dual roles, their stressors, and the level of function required, were complexities not accounted for in the primary analysis. We found that, while women's experience is now perceived as different from men's, this phenomenon is rarely investigated systematically or written into research protocols (Holstein and Harding, 1992). Some exceptions to the rule are the investigations by M. Bleuler (1978), Gilligan (1982), Seeman (1985), and Test and Berlin (1983).

Strategy #2: Comparison Across Similar Studies

The five studies cited earlier in the chapter possess the same goals but use different diagnostic systems for determining schizophrenia, varying lengths of follow-up and methodology, among other factors, which make comparison between them difficult. They give a strong indication of the trend toward recovery only when all five are clumped together because, despite their differences, the similarity of their findings is remarkable (see Table 1). One-half to two-thirds of nearly 1200 patients followed over two to three decades significantly improved or recovered. The reason that the Vermont/Maine comparison is so important is the fact that these are the only two studies in the very long-term literature of schizophrenia and other serious mental illness which were matched in sample, catchment area, treatment eras, diagnostic criteria, design, and methodology (DeSisto et al., submitted, a, submitted, b).

Table 1. Five Recent Long-Term Studies of Schizophrenia

Investigators	Sample size	Average length In years	% Subjects recovered and/or improved significantly
M. Bleuler (1972), Burghölzli - Zurich	208	23	(53 - 66)
Huber et al. (1979), Bonn Studies	502	22	57
Ciompi and Müller (1976), Lausanne investigations	289	37	53
Tsuang et al. (1979), Iowa 500	189	35	46
Harding et al. (1987), Vermont	118	32	62 - 68

Strategy #3: Subgroup Analysis Within the Same Study

A third technique has been to clump life course markers as grouped means within certain subgroup categories (e.g., type of onset, severity, degree of chronicity, diagnosis, age of onset, demographics, etc.). We discovered that with a seemingly simple sub-grouping such as dividing the cohort into age brackets and correlating these groups with outcome measures, we found no significant relationships. The between-subject heterogeneity in other time-related variables (such as age at admission and duration of time since first admission) washed out

the effects of presumed homogeneous subgroupings (Harding et al., 1987c). As an experiment, we then chose every subject with a 42-year course. Wide heterogeneity appeared. The age range for those subjects went from 57 to 84. When we subtracted the 42 years course from each person's age, we ranged from age 15 to age 42 as their ages at first admission. These differences should reflect diverse prognostic indicators. This problem of heterogeneity underlying key measures points to the need for an analysis in which the association of age with outcome is assessed by partialing out the effects of age at first admission and thus length of course (Harding, et al., 1987c). In pursuing this strategy, we discovered that the oldest subjects who had the shortest course (20 - 29 years) achieved the best Global Assessment Scale (Endicott, et al, 1986) scores at outcome. In addition to the importance of these methodologic questions for interpreting results, it should be pointed out that the relevant developmental issues confronting people at various ages of onset have not begun even to be characterized (Harding et al. 1987c). We have wondered if we finally achieve some sense of true homogeneous subgroups, whether our cell sizes will become too small to provide meaningful analyses at all?

Strategy #4: The Individual as His or her Own Control

Should a person be measured against his own baseline? Such a strategy might be a good idea, but engenders the following problems. First, the investigators should know the subjects across the period under scrutiny. In the Vermont Project, we were fortunate that the five clinicians of the original project team were able to maintain contact across 32 years with many members of the cohort and are in a good position to say, for example, that: "Barbara has come a long way. She was once withdrawn and apathetic and now, after several years of improvements and regressions, she consistently goes out once a week in the company of friends through her own initiative." M. Bleuler (1978) is also in a position to make such judgments.

In addition, we came to appreciate that a person's baseline keeps changing as he or she proceeds through developmental tasks. Looking for a return to premorbid status may not be as valid an indicator of recovery as we think it is. Who wants to be judged by their 18-year-old status when they are emerging from an illness at age 35? Further, adult development appears to proceed across domains of function in a fashion close to Piaget's notion of "horizontal decalage" (see Ginsberg and Opper, 1979). (We have re-interpreted his term to describe such natural adult behaviors as the use of *dialectical thinking* in a discussion on ethics; *formal operations* when dealing with the bank; *concreteness* when cooking lunch; and being *pre-operational* when dealing with statistics [all in the same hour]. Are not the varying degrees of functioning within the same person further evidence for Strauss and Carpenter's idea of open-linked systems (Strauss and Carpenter, 1974)? Therefore, given all the underlying complexities, if we measure people against their own baselines, we would probably have 269 single case studies and lose our funding.

THE EFFECT OF THE RATER'S EPISTEMOLOGY

Underlying all efforts is the pervasive problem of recovery measured against the theoretical constructs evolved by the current state of the art and individually interpreted by the investigator. The epistemological framework would define whether recovery meant cured, in remission, or a retreat to underlying vulnerability (such as suggested by Zubin and Spring, 1977). To be cured would mean that a new episode would be considered a new illness. Some cultures believe in this idea very strongly, even about schizophrenia. For example, Waxler (1979) described the Sri Lankan viewpoint as a contributor to a better recovery rate due to the resulting higher expectation of functioning and the lightening of the psychological burden of on-going illness from the patient and his family.

We note that being "in remission" carries with it a heavy impending time bomb effect. Robert Cancro (1982) once commented that women are not considered to be latently pregnant. Further, he suggested that theoretical frameworks made you decide whether a Kansan who moved to Missouri could now be considered a Missourian, or if the person was *born* in Kansas, was he or she a Kansan no matter where he or she went or how long he or she lived elsewhere? Although Cancro was referring to the genetics of schizophrenia, we might refocus the discussion and pose the question, "How long is remission before recovery can be claimed?

We suggest replacing the outlook, which views a person as having a lifetime of illness with intervals of remission included, with the view of a person with a life course of work and relationships, developmental lags and spurts with episodes of illness included. When considering a subject as having a "life", we then must focus on the person behind the disorder (Bleuler, 1978) and not on the illness itself. Paul Lieberman (1984) at Dartmouth has suggested. The same patient is never the same person at each admission". There seems to be a person x illness x environment interaction which continues to reshape the on-going process (Harding and Strauss, 1985; Strauss, et al. 1985).

DOUBLE STANDARDS IN JUDGING RECOVERY

Participation in case conferences and review of instrument batteries has led us to wonder if we, as clinicians and investigators, tend to have a double set of standards in judging how well someone is functioning? We all seem to have a certain but nebulous sense of what is "normal". I wonder if our standard for patients is not "super normal" (e.g., as in Maslow's (1954) self-actualized persons, few and famous? For example, the top rating for the Global Assessment Scale (91-100) uses the following criteria: "No symptoms, superior functioning in a wide range of activities, life's problems never seem to get out of hand, and is sought out by others because of his warmth and integrity" (Endicott et al., 1976).

Quite possibly we secretly harbor the idea that maybe everyone else we know is happier, handles things better, does not worry about the mortgage coming due, spreading waistlines, gray hair, or falling on one's face during an important presentation. When we do discover that other people worry too, get depressed, stumble and fall, only to laugh again, it is a celebration and quite often the glue that binds a friendship. For ourselves and our friends we may tend to accept the sets of idiosyncrasies and horizontal decalage because we know where we have come from and the direction in which we are going with a thorough-going appreciation of underlying continuities. Do these understandings permeate clinicians' judgments and investigators' ratings often, or do patients have to perform twice as well to be considered half as good?

SUMMARY

In summary, this chapter examined some of the complexities surrounding the measurement of recovery from severe psychiatric disorder. Investigators must be careful in employing the words "outcome" or "end state" because they end up being interpreted by clinicians and clients alike as a real phenomenon. It has been suggested that we substitute outcome with the phrase "markers of course" or simply "at this point in the life course." Longitudinal documentation comes closer to recapturing course of life than do two cross-sectional assessments.

Questions about the measurement and meaning of recovery can best be answered by assessment of multiple domains across long periods of time. The problems in declaring recovery begin with the realization that human beings develop differentially across domains of functioning and across time. Four strategies for assessment have been discussed: 1) a person might be measured against his or her own baseline; 2) the employment of grouped means within sub-samples; 3) the possibility of comparison of groups across studies; and 4) measurement

against some construct of normal behavior or outcome. The advantages and disadvantages of each approach were presented.

A philosophical look at the impact of epistemology upon one's view of recovery as cure, remission, or underlying vulnerability was discussed as well as a postulated double standard employed by clinicians for rating recovery in patients versus rating functioning in friends. We joke about the fact that no one we know as colleagues, including ourselves, could meet the GAS score of 91 to 100, yet we persisted in using this scale and others as earnest assessments of subjects. It should be noted that the Global Assessment of Functioning (GAF) which recently replaced the GAS has dropped the 91-100 rating (APA, 1987, pp 12 and 20).

In conclusion, I would venture to say that to designate someone as recovered from severe psychiatric disorder is a judgment call as challenging as the decision made about when to call behavior an illness. It is trying to pin down a constantly moving target. In addition, the process encompasses cultural expectations, the state of the art, and the personality of the investigator, all of which influence the theoretical framework in which he or she constructs and selects the questions for the assessment, the manner in which they are asked, and how the data are analyzed.

ACKNOWLEDGEMENTS

This chapter is a revision of an article which appeared in the Psychiatric Journal of the University of Ottawa, 11 (4),1986 and is reproduced by permission of Yvon La Pierre, M.D., Chairman, and the Journal Management Committee, Department of Psychiatry at the University of Ottawa. The article was first a banquet presentation for G.W.A.N. (The International Psychiatric Research Society), given on November 23, 1985, in Toronto, Ontario. This effort was also supported in part by NIMH Grants R 01 MH29575 and R 01 MH40607. The ideas contained in this chapter have been sifted from hundreds of conversations across the last 15 years within the Vermont and Maine Project Teams as well as with friends and colleagues ... anyone who would sit still long enough to discuss the issues.

REFERENCES

American Psychiatric Assn. (1980). Diagnostic and Statistical Manual, 111, APA Press, Washington, DC.

American Psychiatric Assn. (1987). Diagnostic and Statistical Manual, III-R, APA Press, Washington, DC.

Bland, R.C. and Orn, H. (1980). "Schizophrenia: Schneider's first-rank symptoms and outcome", Br. J. Psychiatry, 137, 63 - 68.

Bleuler, E. (1950). Dementia Praecox or the Group of Schizophrenias (translated by J. Zinkin), International Universities Press, New York.

Bleuler, M. (1978). Die schizophrenen Geistesst6rungen im Lichte Kranken - und Familiengeschiten. George Thieme Verlag, Stuttgart (1972) (Translated by S.M. Clemens, The Schizophrenic Disorders, Long-term Patient and Family Studies, Yale University Press, New Haven.

Cancro, R. (1982). "The role of genetic factors in the etiology of schizophrenic disorders", The Schizophrenic Disorders. Part 11 of the Annual Review, 135th annual meeting, American Psychiatric Association.

Chiftick, R.A., Brooks, G.W., Irons, F.S. and Deane, W.N. (1961). The Vermont Story, Queen City Printers, Burlington, VT.

Ciompi, L. (1980). "Catamnestic long-term study on the course of life and aging in schizophrenia", Schizophr. Bull, 6(4), 592 - 605.

Ciompi, L. and Moller, C. (1976). Lebensweg und Alter Schizophrenen. Eine katamnestic Lonzeitstudies bis ins senum, Springer - Veriag, Berlin.

Consalvo, C.M., Harding, C.M., Landerl, P.D. and Mikkelsen, J., (1981). The Community Assessment Scale (CAS), Waterbury, VT.

Deane, W.N., Brooks, G.W. (1967). Five-year follow-up of chronic hospitalized patients, Report from Vermont State Hospital, Waterbury, VT.

DeSisto, M., Harding, C.M., Ashikaga, T., McCormick, R. and Brooks, G.W. (submitted, a). "The Maine - Vermont Project: a matched cohort comparison of long-term outcome 1. Methodology, study samples, design, overall 36-year outcome for the Maine cohort."

DeSisto, M., Harding, C.M., Ashikaga, T., McCormick, R. and Brooks, G.W.(submitted, b) "The Maine - Vermont Project: a matched cohort comparison of long-term outcome II. Cross-sectional outcome for both cohorts".

Endicott, J., Spitzer, R.L., Fleiss, J.L. and Cohen, R. (1976). "The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance", Arch. Gen. Psychiatry, 33, 766 - 771.

Gardos, G., Cole, J.O. and LaBrie, R.A. (1982). "A 12-year follow-up study of chronic schizophrenics", Hosp. Community Psychiatry, 33(12), 983 - 984.

Gilligan, C. (1982). In a Different Voice: Psychological Theory and Women's Development, Harvard University Press, Cambridge, MA.

Ginsberg, H. and Opper, S. (1979). Piaget's Theory of Intellectual Development,2nd edn, Prentice - Hall, EnglewoodCliffs, pp. 152,109,159.

Harding, C.M. and Strauss, J.S. (1985). "The course of schizophrenia: an evolving concept". In Controversies in Schizophrenia: Changes and Constancies (ed. M. Alpert), Guilford Press, New York, pp. 339 - 353.

Harding, C.M., Brooks, G.W., Ashikaga, T., Strauss, J.S. and Breier, A. (1987a). "The Vermont longitudinal study of persons with severe mental illness: 1. Methodology, study sample, and overall status 32 years later", Am. J. Psychiatry, 144(6), 718 - 726.

Harding, C.M., Brooks, G.W., Ashikaga, T., Strauss, J.S., Breier, A. (1987b). "The Vermont longitudinal study: II. Long-term outcome of subjects who retrospectively met the criteria for DSM-111 schizophrenia", Am. J. Psychiatry, 144(6), 727 - 735.

Harding, C.M., Brooks, G.W., Ashikaga, T., Strauss, J.S. and Landerl. P.D. "Aging and social functioning in once-chronic schizophrenic patients 22 - 62 years after first admission: The Vermont Story". In Schizophrenia, Paranoia and Schizophreniform Disorders in Later Life (eds N. Miller and G.D. Cohen), Guilford Press, New York, pp. 160 - 166.

Harris, A., Linker, I., Norris, V. and Shepherd, M. (1956). "Schizophrenia: a prognostic and social study", Br. J. Prev.Soc. Med., 10, 107.

Hawk, A.B., Carpenter, W.T. and Strauss, J.S. (1975). "Diagnostic criteria and 5-year outcome in schizophrenia: a report from the International Pilot Study of Schizophrenia", Arch. Gen. Psychiatrv, 32, 343 - 347.

Hogarty, G.E. and Katz, M.M. (1971). "Norms of adjustment and social behavior", Arch. Gen. Psychiatry, 25(11), 470 - 480.

Holmbe, R. and Astrup, A. (1957). "A follow-up study of 255 patients with acute schizophrenia and schizophreniform psychosis", Acta Psychiatr. Scand, Suppl. 115.

Holstein, A.R. and Harding, C.M. (1992). "Omissions in assessment of work roles: implications for evaluating social functioning and mental health", J. Orthopsychiatry, 62(3), 469 - 474.

Holt, W.L., Jr and Holt, W.M. (1952). "Long-term prognosis in mental illness: a thirty-year follow-up of 141 mental patients", Am. J. Psychiatry, 108, 735 - 739.

Huber, G., Gross, G. and SchOttler, R. (1979). Schizophrenie: Verlaufs und sozialpsychiatrische Lanqzeitundersuchungen an den 1945 bis 1959 in Bonn Hospitalisierten schizophrenen Kranken. Monographien aus dem Gesamtgebiete der Psychiatrie. Bd. 21., Springer - Verlag, Berlin.

Leighton, A.H. and Leighton, D.P. (1949). "Gregorio, the handtrembler: a psychobiological personality study of a Navaho Indian", Peabody Museum Papers, Harvard University Press, Cambridge, MA.

Lieberman, P. (1984). Personal Communication.

Lo, W.H. and Lo, T. (1977). "A ten-year follow-up study of Chinese schizophrenics in Hong Kong", Br. J. Psychiatry, 131, 63 - 66.

Maslow, A. (1954). Motivation and Personality. Harper & Row, New York.

Meyer, A. (1919). "The life chart", In Contributions to Medical and Biological Research. Paul B. Hoeber, New York.

Overall, J.E. and Gorham, D.R. (1962). Brief Psychiatric Rating Scale (BPRS). Psychol. Rep., 10, 799 - 812.

Rennie, T.A.C. (1939). "Follow-up study of five hundred patients with schizophrenia admitted to the hospital", Arch. Neurol. Psychiatry, 42(5), 877 - 891.

Seeman, M.V. (1985). "Sex and schizophrenia", Can. J. Psychiatry, 30(8), 313 - 315.

Stephens, J.H. (1970). "Long-term course and prognosis in schizophrenia", Semin. Psychiatry, 2(4), 464 - 485.

Strauss, J.S. and Carpenter, W.T. (1974). "Prediction of outcome in schizophrenia. II. Relationships between predictor and outcome variables", Arch. Gen. Psychiatry, 31, 37 - 42.

Strauss, J.S. and Harder, D.W. (1981). "The case record rating scale: a method for rating symptom and social function data from case records", Psychiatry Res, 4, 333 - 345.

Strauss, J.S., Kokes, R.F., Ritzler, B., Harder, D.W. and VanOrd, A. (1978). "Patterns of disorder in first admission psychiatric patients", J. Nerv. Ment. Dis., 166(9), 611 - 625.

Strauss, J.S., Hafez, H., Lieberman, P. and Harding, C.M. (1985). "The course of psychiatric disorder: longitudinal principles", Am. J. Psychiatry, 142(3), 289 - 296.

Test, M.A. and Berlin, S.B. (1981). "Issues of special concern to chronically mentally ill women", Professional Psychol., 12(1), 136 - 145.

Tsuang, M.T., Woolson, R.F. and Fleming, J.A. (1979). "Long-term outcome of major psychoses. 1. Schizophrenia and affective disorders compared with psychiatrically symptom-free surgical conditions", Arch. Gen. Psychiatry, 36, 1295 - 131.

Vaillant, G.E. (1975). Paper presented at the 128th annual meeting of the American Psychiatric Association, Anaheim, CA.

Waxler, N. (1979). "Is outcome for schizophrenia better in non-industrial societies? The case of Sri Lanka", J. Nerv. Ment. Dis., 67, 144 - 158.

World Health Organization (1973). The International Pilot Study of Schizophrenia, WHO Press, Geneva.

World Health Organization (1979). Schizophrenia: An International Follow-up Study, Wiley, Chichester.

Zubin, J. and Spring, B. (1977). "Vulnerability: a new view of schizophrenia", J. Abnorm. Psychology, 86, 103 - 126.